Molecular acidity: A quantitative conceptual density functional theory description

Shubin Liu,[1,a] Cynthia K. Schauer,[2] and Lee G. Pedersen[2,3]
[1]*Research Computing Center, University of North Carolina, Chapel Hill, North Carolina 27599-3420, USA*
[2]*Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290, USA*
[3]*Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, P.O. Box 12233, North Carolina 27709, USA*

(Received 19 July 2009; accepted 30 September 2009; published online 23 October 2009)

Accurate predictions of molecular acidity using *ab initio* and density functional approaches are still a daunting task. Using electronic and reactivity properties, one can quantitatively estimate pKa values of acids. In a recent paper [S. B. Liu and L. G. Pedersen, J. Phys. Chem. A **113**, 3648 (2009)], we employed the molecular electrostatic potential (MEP) on the nucleus and the sum of valence natural atomic orbital (NAO) energies for the purpose. In this work, we reformulate these relationships on the basis of conceptual density functional theory and compare the results with those from the thermodynamic cycle method. We show that MEP and NAO properties of the dissociating proton of an acid should satisfy the same relationships with experimental pKa data. We employ 27 main groups and first to third row transition metal-water complexes as illustrative examples to numerically verify the validity of these strong linear correlations. Results also show that the accuracy of our approach and that of the conventional method through the thermodynamic cycle are statistically similar. © *2009 American Institute of Physics*. [doi:10.1063/1.3251124]

I. INTRODUCTION

Molecular acidity, the ability or tendency of an acid to lose a proton, is one of the fundamental and important physiochemical properties of a molecule, essential in many chemical processes such as enzymatic reactions and drug metabolisms. Since the process takes place in aqueous conditions, computational simulation using *ab initio* or density functional theory (DFT),[1–7] or through statistical or empirical force field approaches (such as molecular mechanics and quantitative structure-reactivity relationships),[8,9] is not at all trivial. Recent efforts in literature attempt to make use of global descriptors such as frontier molecular orbitals[7,9–13] to predict this quantity have met with limited success.

Under the hypothesis that molecular acidity is a local property of a molecule limited to the neighboring region of the particular acidic atom only and that the impact of the environment (such as substituent groups, solvent effect, etc.) is reflected through the changes in electronic or stereoelectronic properties to that small neighborhood, we recently proposed to quantitatively predict molecular acidity through two equivalent quantum descriptors, molecular electrostatic potential (MEP) on the acidic atom[14] (i.e., O for carboxyl acid and alcohols, N for amines and anilines, and S for sulfonic acids and thiols) and the sum of the valence natural atomic orbitals (NAOs) of the same atom. Strong linear correlations between experimental pKa values and these quantities have been disclosed. Moreover, if the MEP is subtracted by the isolated atomic MEP, we observed a single unique linear relationship between the resultant MEP difference and experimental pKa data of amines, anilines, carbonyl acids, alcohols, sulfonic acids, thiols, and their substituents. These results can be utilized to simultaneously estimate pKa values at multiple sites of a molecule with a single DFT or *ab initio* calculation in gas phase.

In this work, we extended our early work by establishing molecular acidity based on the foundation of conceptual DFT.[15–18] To that end, we will show that there exist similar linear relationships for the dissociating proton and that these correlations are the natural consequence of the Taylor expansion of the total energy of the system from conceptual DFT at the first-order approximation. We demonstrate the validity of these arguments with numerical data from 27 main groups and first, second, and third row transition metal-water complexes.

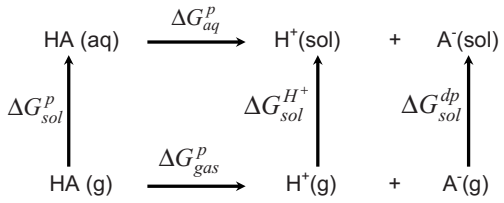
II. THEORETICAL FRAMEWORK

In theory, the pKa value of an acid, HA, is defined as the negative logarithm, $\mathrm{pKa} = -\log_{10} \mathrm{Ka}$, of the equilibrium constant Ka of the following acid dissociation:

$$\mathrm{HA} \rightleftharpoons \mathrm{A}^- + \mathrm{H}^+, \quad \mathrm{Ka} = [\mathrm{HA}]/[\mathrm{A}^-][\mathrm{H}^+], \tag{1}$$

where HA is a generic acid which dissociates by splitting into the conjugate base A^- of the acid, and a proton, H^+, in the aqueous condition, and $[HA]$, $[A^-]$ and $[H^+]$ are the equilibrium concentration of these species, respectively. In thermodynamics, an equilibrium constant Ka is related to the standard Gibbs free energy change for the above dissociation reaction, so for an acid dissociation constant,

[a]Author to whom correspondence should be addressed. Electronic mail: shubin@email.unc.edu.



SCHEME 1.

$$\Delta G^\circ = 2.303RT\mathrm{pKa}, \tag{2}$$

where R is the gas constant and T is the temperature in Kelvin.

In practice, the pKa value can be computed through the following thermodynamic cycle (Scheme 1) using *ab initio* and DFT methods:[1,19]

$$2.303RT\mathrm{pKa} = \Delta G^p_{\mathrm{aq}} = \Delta G^{dp}_{\mathrm{sol}} + \Delta G^{\mathrm{H}^+}_{\mathrm{sol}} - \Delta G^p_{\mathrm{sol}} + \Delta G^p_{\mathrm{gas}}, \tag{3}$$

where R is the Rydberg gas constant and T is the temperature. ΔG^p_{aq} is the sum of the free energy of deprotonation of the gas-phase species $\Delta G^p_{\mathrm{gas}}$, the free energies of desolvation of the protonated form $-\Delta G^p_{\mathrm{sol}}$ and solvation of the deprotonated form $\Delta G^{dp}_{\mathrm{sol}}$, and the free energy of solvation for the proton $\Delta G^{\mathrm{H}^+}_{\mathrm{sol}}$. Even though tremendous progress has been made in recent decades in accurate prediction of pKa values,[1–9] for medium to larger-sized systems, these calculations are still computationally demanding. Besides, the computational accuracy of pKa values for certain systems was found to be unfaithful even with advanced methodologies.[7]

Recently, we proposed to use the MEP at the nucleus of the acidic atom (e.g., O for carboxyl acids and alcohols, N for amines, S for sulfonic acids and thiols, etc.),[14]

$$V_{\mathbf{R}_A} = \sum_{i \neq A} \frac{Z_i}{|\mathbf{R}_i - \mathbf{R}_\mathbf{A}|} - \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_\mathbf{A}|} d\tau, \tag{4}$$

to quantitatively estimate molecular acidity and also discovered another equally reliable predictor, the sum of the valence NAO energies of the acidic atom. These two descriptors were found to be closed related, with the correlation coefficient R^2 of the two around 0.99 for a large sample of data sets. No dependence of these relationships on methods (Hartree–Fock or DFT) or basis sets has been verified.[14]

Molecular acidity has also been of considerable study in the field of so-called conceptual DFT,[15–18] also called density functional reactivity theory. It concerns with the scenario when changes in the number of electrons (ΔN) and the external potential ($\Delta v(\mathbf{r})$) take place, subsequent variations in the total energy of the system, $E[v(\mathbf{r});N]$, which is the state function of the total number of electrons N and the external potential $v(\mathbf{r})$, will follow, obeying the following relationship:

$$\begin{aligned}\Delta E &= E[v + \Delta v; N + \Delta N] - E[v;N] \\ &= \left\{ \left(\frac{\partial E}{\partial N}\right)_{v(\mathbf{r})} \Delta N + \int \left(\frac{\delta E}{\delta v(\mathbf{r})}\right)_N \Delta v(\mathbf{r}) d\mathbf{r} \right\} \\ &\quad + \frac{1}{2}\left[\left(\frac{\partial^2 E}{\partial N^2}\right)(\Delta N)^2 + 2\Delta N \int \frac{\delta\, \partial E}{\delta v(\mathbf{r})\, \partial N} \Delta v(\mathbf{r}) d\mathbf{r} \right. \\ &\quad \left. + \int\int \left(\frac{\delta^2 E}{\delta v(\mathbf{r}') \delta v(\mathbf{r})}\right)_N \Delta v(\mathbf{r}) \Delta v(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \right] \\ &\quad + (\text{third and higher order terms}).\end{aligned} \tag{5}$$

The terms in braces, { }, are the first-order terms and the terms in brackets, [ ], are the second-order contributions. Third and higher order terms are usually assumed to be qualitatively unimportant and thus can be omitted.

The process of acid dissociation, Eq. (1), is a special case of the above general consideration. In Eq. (1), one proton is dissociated from the acid molecule, leading to the change in external potential $\Delta v(\mathbf{r})$, but the total energy of electrons in the acid system remains unchanged. With $\Delta N = 0$, up to the second order, Eq. (5) is simplified to be

$$\begin{aligned}\Delta E[\Delta v(\mathbf{r})] &= \int \rho(\mathbf{r}) \Delta v(\mathbf{r}) d\mathbf{r} \\ &\quad + \frac{1}{2}\int\int \chi(\mathbf{r},\mathbf{r}') \Delta v(\mathbf{r}) \Delta v(\mathbf{r}') d\mathbf{r} d\mathbf{r}',\end{aligned} \tag{6}$$

where we made use of the fact that the electron density of the system $\rho(\mathbf{r}) = (\delta E / \delta v(\mathbf{r}))_N$ and $\chi(\mathbf{r},\mathbf{r}') = (\delta^2 E / \delta v(\mathbf{r}) \delta v(\mathbf{r}'))_N = (\delta \rho(\mathbf{r}) / \delta v(\mathbf{r}'))_N$ as the linear-response function of the electron density. Proposals of how to obtain optimal change in $\Delta v(\mathbf{r})$ and subsequently ΔE have recently been investigated, leading to the novel concepts such as potentialphilicity, potentialphobicity, chargephilicity, and chargephobicity.[20,21]

Furthermore, in the simplest case when only the first-order term is considered, Eq. (6) becomes

$$\Delta E[\Delta \nu(\mathbf{r})] = \int \rho(\mathbf{r}) \Delta \nu(\mathbf{r}) d\mathbf{r}. \tag{7}$$

Now, for the special case of Eq. (1), the change in the external potential, $\Delta v(\mathbf{r})$, resulted from the dissociation of a proton, can explicitly be obtained, giving

$$\Delta \nu(\mathbf{r}) = \sum_{i \neq H} \frac{Z_i}{|\mathbf{R}_i - \mathbf{R}_H|} - \frac{1}{|\mathbf{r} - \mathbf{R}_H|}, \tag{8}$$

where $\mathbf{R}_H$ is the coordinate of the leaving proton and $\{\mathbf{R}_i\}$ are the coordinates of the other nuclei in the acid molecule. The first term in Eq. (8) is the nuclear-nuclear repulsion potential between the leaving proton and the other nuclei in the system, and the second term is the attraction potential between an electron at position $\mathbf{r}$ and the leaving proton at $\mathbf{R}_H$.

Substitution of Eq. (8) into Eq. (7) leads to

$$\Delta E[\Delta \nu(\mathbf{r})] = \sum_{i \neq H} \frac{Z_i}{|\mathbf{R}_i - \mathbf{R}_H|} - \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_H|} d\tau. \tag{9}$$

Thus, the change in the total electronic energy of an acid due to dissociation of a proton equals the electrostatic potential

on the same proton. As the main result of the present work, this formula suggests that the MEP of the leaving proton can serve as a linear predictor of the pKa of the acid,

$$\mathrm{pKa} \propto \sum_{i \neq H} \frac{Z_i}{|\mathbf{R}_i - \mathbf{R}_H|} - \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_H|} d\tau. \quad (10)$$

At least three approximations have been employed to obtain Eq. (10). First, we neglected thermodynamic contributions from solvent and temperature effects and approximated the Gibbs free energy change in Eq. (2) by the total electronic energy difference in Eq. (9). Second, the Taylor expansion in Eq. (5) was truncated at first order. Third, we assumed in Eq. (8) that before and after the proton dissociation, there is an undetectable structure relaxation in the conjugate base A^- so the only change in external potential is due to the removal of the proton from the acid.

Notice that Eqs. (4) and (9) have the same format but Eq. (4) is for the acidic atom in an acid such as O in carboxyl acid and alcohols, N in amines and anilines, and S in sulfonic acids and thiols, whereas Eq. (9) is the nuclear MEP of the leaving proton. In this work, using 27 metal-water complexes as an example, we examine the validity of Eq. (10) and at the same time verify the effectiveness of our early results to estimate pKa values with MEP at the nucleus of acidic atoms and the sum of natural valence atomic orbital energies.

III. COMPUTATIONAL DETAILS

A total of 27 metal-water complexes, $M(H_2O)_n^{m+}$ (M =Al, Ba, Be, Ca, Co, Cr, Fe, Ga, Hf, In, K, Li, Mg, Mn, Na, Ni, Sc, Sr, Ti, Tl, Zn, and Zr; m=1,2,3,4; n=4,6), range from main group to first, second, and third row transition metal cations with different positive charges and different coordination numbers have been investigated in this work. The positive charge on the metal cation ranges from +1 (Li, K, Na, and Tl) to +4 (Hf, Sn, Ti, and Zr) and their coordination number of water molecules is 6 except for $Be(H_2O)_4^{2+}$ and $Tl(H_2O)_4^{3+}$. Experimental pKa values of these species are from literature,[22] the range of which spans about 19 orders of magnitudes, between −4.0 in $Ti(H_2O)_6^{4+}$ and +14.7 in $Na(H_2O)_6^{+}$. To simulate the first-shell solvent effect, an extra 12 water molecules were added outside each complex, with each original water molecule hydrogen bonded to two outside water molecules, which can be denoted by $M(H_2O)_n^{m+}(H_2O)_{12}$. The ONIOM (Our own N-layered Integrated molecular Orbital and molecular Mechanics) model has been employed to investigate the impact from the second-shell solvent effect, for which the theory level of RHF/6-31G(d) was employed for the middle layer and semiempirical AM1 approach was used for the low layer, whereas the high layer is the $M(H_2O)_n^{m+}$, treated at the same level of theory as the gas phase. As an illustrative example, Fig. 1 displays the optimized structure of the $Al(H_2O)_6^{3+}$ complex in gas phase [Fig. 1(a)], in the condensed phase [Fig. 1(b)] simulated by 12 additional water molecules to model the first shell of solvent molecules through the explicit solvent model, and in the condensed phase [Fig. 1(c)] simulated through the three-layer ONIOM model by adding 30 additional water molecules as the first shell (middle layer of ONIOM) of solvent molecules and 64 additional water molecules to model the second shell (low layer of ONIOM) of solvent molecules. A full structure optimization for all the models was first carried out at the DFTB3LYP/6-311+G(d) (Refs. 23–25) level except for the metal elements whose above triple-zeta basis set is unavailable. In those cases, we used the LanL2DZ basis set for them instead.[26] Other higher level basis sets were tested and no qualitatively different result was observed. When a molecule had multiple spin states,[27–30] all were examined and the one with the lowest energy was picked for the subsequent study. After the structure optimization, a frequency calculation for each of the systems was carried out to check that the optimized structure was indeed a minimum (i.e., with no imaginary frequency). Then, single point calculations were performed to obtain the MEP on each of the nuclei followed by a full natural bond orbital (NBO) analysis[31] to obtain the NAO energies. To compare our method with the results obtained from the thermodynamic cycle[1] (Scheme 1), the structure of an acid and its conjugate base in both gas and aqueous phases has been optimized at the same level of theory and then a frequency calculation is followed to obtain their free energy. We employed the implicit solvent polarizable continuum model, CPCM (Conductor-like Polarizable Continuum Model),[32,33] to simulate the solvent effect in the aqueous phase calculations. The free energy values for the proton in both phases, $G(H^+_{gas})$ and $\Delta G^{H^+}_{sol}$ are derived from experiment. We have used the values $G(H^+_{gas})$ =−6.28 kcal/mol and $\Delta G^{H^+}_{sol}$=−264.61 kcal/mol.[34,35] All calculations have been performed with the GAUSSIAN 03 package[36] with tight self-consistent field convergence and ultrafine integration grids.

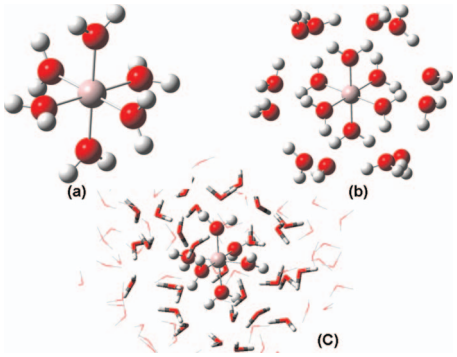


FIG. 1. Three models of the optimized structure for the $Al(H_2O)_6^{3+}$ complex: (a) in the gas phase, (b) in the condensed phase simulated by 12 additional water molecules to model the first shell of solvent molecules through the explicit solvent model, and (c) in the condensed phase simulated through three-layer ONIOM model by adding 30 additional water molecules as the first shell [middle layer of ONIOM at the level of RHF/6-31G(d) theory] of solvent molecules and 64 additional water molecules to model the second shell (low layer of ONIOM using semiempirical AM1 approach) of molecules. The high ONIOM layer employs the DFTB3LYP/6-311+G(d) approach.

IV. RESULTS AND DISCUSSION

Table I shows the results according to our earlier work,[14] where MEP on the oxygen nucleus (MEP_O) and the sum of

TABLE I. Experimental pKa values and computational data for the MEP, HOMO/LUMO, and the sum of natural valence atomic orbital energies of oxygen for 27 metal cation and water complexes. Units in a.u.

Molecules	pKa	2S+1	MEP_O	HOMO	LUMO	ΣNAO_O
$Al(H_2O)_6^{3+}$	5.0	1	−21.803	−0.876	−0.492	−2.639
$Ba(H_2O)_6^{2+}$	13.5	1	−22.073	−0.610	−0.293	−1.801
$Be(H_2O)_4^{2+}$	6.2	1	−21.932	−0.743	−0.356	−2.244
$Ca(H_2O)_6^{2+}$	12.6	1	−22.048	−0.636	−0.299	−1.877
$Co(H_2O)_6^{2+}$	9.6	4	−22.025	−0.639	−0.316	−1.941
$Cr(H_2O)_6^{3+}$	4.0	3	−21.805	−0.847	−0.643	−2.606
$Cr(H_2O)_6^{2+}$	10.0	5	−22.055	−0.518	−0.369	−1.847
$Fe(H_2O)_6^{3+}$	2.0	6	−21.811	−0.866	−0.488	−2.571
$Fe(H_2O)_6^{2+}$	9.5	5	−22.024	−0.612	−0.315	−1.946
$Ga(H_2O)_6^{3+}$	2.6	1	−21.806	−0.872	−0.514	−2.624
$Hf(H_2O)_6^{4+}$	0.2	1	−21.623	−1.052	−0.739	−3.181
$In(H_2O)_6^{3+}$	4.0	1	−21.827	−0.853	−0.490	−2.558
$K(H_2O)_6^{+}$	14.5	1	−22.242	−0.442	−0.151	−1.277
$Li(H_2O)_6^{+}$	13.9	1	−22.228	−0.456	−0.155	−1.318
$Mg(H_2O)_6^{2+}$	11.4	1	−22.028	−0.655	−0.310	−1.936
$Mn(H_2O)_6^{2+}$	10.6	6	−22.032	−0.589	−0.314	−1.922
$Na(H_2O)_6^{+}$	14.7	1	−22.236	−0.449	−0.151	−1.292
$Ni(H_2O)_6^{2+}$	9.9	3	−22.022	−0.649	−0.316	−1.950
$Sc(H_2O)_6^{3+}$	4.3	1	−21.827	−0.852	−0.581	−2.557
$Sn(H_2O)_6^{4+}$	−0.6	1	−21.607	−1.067	−0.711	−3.228
$Sr(H_2O)_6^{2+}$	13.3	1	−22.060	−0.624	−0.297	−1.841
$Ti(H_2O)_6^{3+}$	2.2	2	−21.817	−0.757	−0.608	−2.578
$Ti(H_2O)_6^{4+}$	−4.0	1	−21.593	−1.074	−0.888	−3.242
$Tl(H_2O)_6^{+}$	13.2	1	−22.244	−0.372	−0.155	−1.273
$Tl(H_2O)_4^{3+}$	0.6	1	−21.766	−0.901	−0.583	−2.724
$Zn(H_2O)_6^{2+}$	9.5	1	−22.028	−0.654	−0.317	−1.939
$Zr(H_2O)_6^{4+}$	−0.3	1	−21.627	−1.046	−0.766	−3.162

the three valence 2p NAO energies of the same oxygen atom are exhibited, together with highest occupied molecular orbital (HOMO)/lowest unoccupied molecular orbital (LUMO), spin multiplicity, and experimental pKa values of the 27 metal-water complexes. Linear relationships that stemmed from these quantities are shown in Fig. 2, where one finds strong correlations between MEP_O and experimental pKa values [Fig. 2(a), $R^2=0.932$], between the sum of the three 2p NAO energies and MEP_O [Fig. 2(b), $R^2=1.000$], and between chemical potential and experimental pKa data as well [Fig. 2(c), $R^2=0.922$]. According to Mulliken,[15] chemical potential is defined as $\mu \equiv \frac{1}{2}(I+A)$ with I and A as the first ionization energy and electron affinity, respectively, which can be approximated by the HOMO and LUMO energies via $I \approx -\varepsilon_{HOMO}$ and $A \approx -\varepsilon_{LUMO}$. We also considered other global reactivity indices from conceptual DFT such as hardness[37–40] and electrophilicity index[41–43] but no statistically strong correlation with experimental pKa data was identified. Also, for the explicitly solvated model, $M(H_2O)_n^{m+}(H_2O)_{12}$ [Fig. 1(b)], and the explicit solvent model from the ONIOM method [Fig. 1(c)], similar relationships among these quantities have also been observed. For example, the linear correlation coefficients, R^2, between MEP and experimental pKa values for the first-shell solvent model [Fig. 1(b)] and second-shell solvent mode [Fig. 1(c)] are 0.924 and 0.919, respectively. These results, as already have been demonstrated earlier,[14] indicate that the addition of the solvent effect does not invalidate the simpler model.

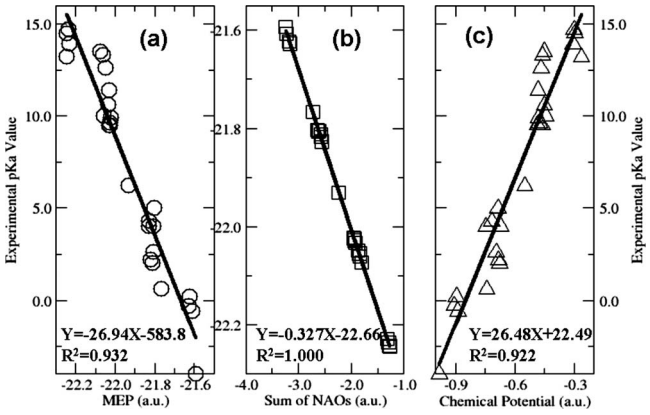


FIG. 2. Linear relationships between MEP at oxygen nucleus (O MEP), the sum of valence NAO energies of oxygen, experimental pKa values, and chemical potential derived from HOMO and LUMO energies for the 27 $M(H_2O)_n^{m+}$ complexes.

The correlation of experimental pKa data to computed pKas for the thermodynamic cycle calculations (using Scheme 1) is shown in Fig. 3(a). A similar plot [Fig. 3(b)] shows the correlation for our MEP fitting procedure. The dashed lines in Fig. 3 represent the ideal state of perfect agreement between experiment and computed pKas. As seen from Fig. 3(a), for some compounds whose pKa values are large such as $Ca(H_2O)_6^{2+}$, $Fe(H_2O)_6^{2+}$, $Ni(H_2O)_6^{2+}$, and $Zn(H_2O)_6^{2+}$, the computed and experimental values are very close. However in other cases, especially for complexes with

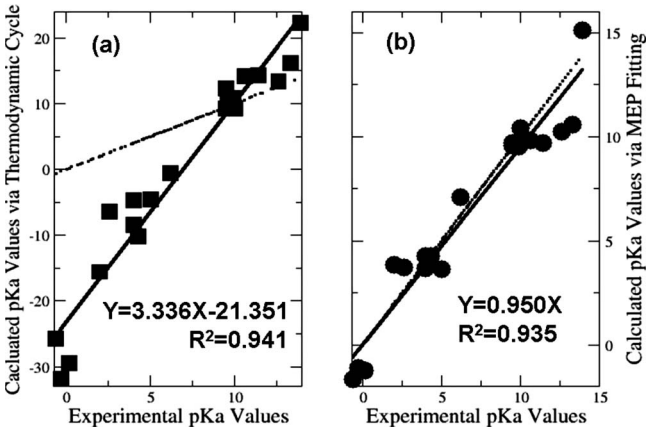


FIG. 3. Comparison of experimental pKa data with (a) calculated pKa values obtained from the thermodynamic cycle in Scheme 1 and (b) calculated pKa values through MEP fitting from this work. The dashed lines represent the ideal state of perfect agreement between the experimental and computed pKas.

small pKa values such as $Fe(H_2O)_6^{3+}$, $Sn(H_2O)_6^{4+}$, $Ti(H_2O)_6^{4+}$, and $Zr(H_2O)_6^{4+}$, we see substantial discrepancy between the two. The different behaviors of the metal-water complexes with different pKa values likely originate from the different performances of the CPCM implicit solvent model at the different pH range. Improved pKa predictions can be obtained using the thermodynamic cycle approach if empirical scaling parameters and procedures are introduced.[44] Overall, for the set of molecules in this study, the correlation between experimental pKa data and the calculated values from thermodynamic cycle method is strong, with the correlation coefficient $R^2=0.94$. The accuracy of this method is similar to that of our approach using MEP or NAO whose $R^2=0.94$ as well [Fig. 3(b)] although our approach is computationally less demanding since only one gas-phase calculation is needed.

To verify the validity of Eq. (10), which suggests that as a first-order approximation the pKa value of an acid should be proportional to the electrostatic potential on the nucleus of the leaving proton, shown in Table II are the MEP data at the leaving hydrogen nucleus, MEP_H, as well as the valence (1s) NAO energy of the same atom, $NAO_{H\ 1s}$, in both gas phase and the simulated condensed phase and the experimental pKa values for the 27 metal-water complexes. Strong linear correlations among these data in the gas phase are displayed in Fig. 4, where the correlation coefficients R^2 between MEP_H and pKa [Fig. 4(a)], between $NAO_{H\ 1s}$ and MEP_H [Fig. 4(b)], and between NAO_H and pKa [Fig. 4(c)] data are 0.930, 0.999, and 0.928, respectively. Similar relationships were observed for results in the condensed phase (not shown). Figure 4(a) confirms the validity of Eq. (10), indicating that as was the case for MEP_O, the quantity of MEP_H is a reliable descriptor of the molecular acidity, and corrobo-

TABLE II. The MEP at the hydrogen nucleus (MEP_H) and valence NAO energy of the hydrogen atom ($NAO_{H\ 1s}$) in both gas and condensed phases of 27 main groups and transition metal-water complexes. Units in a.u.

		Gas phase		Condensed phase	
Molecules	pKa	MEP_H	$NAO_{H\ 1s}$	MEP_H	$NAO_{H\ 1s}$
$Al(H_2O)_6^{3+}$	5.0	−0.443	−0.347	−0.631	−0.200
$Ba(H_2O)_6^{2+}$	13.5	−0.701	−0.131	−0.800	−0.048
$Be(H_2O)_4^{2+}$	6.2	−0.569	−0.228	−0.784	−0.054
$Ca(H_2O)_6^{2+}$	12.6	−0.677	−0.147	−0.793	−0.053
$Co(H_2O)_6^{2+}$	9.6	−0.658	−0.158	−0.790	−0.050
$Cr(H_2O)_6^{3+}$	4.0	−0.447	−0.349	−0.632	−0.207
$Cr(H_2O)_6^{2+}$	10.0	−0.649	−0.166	−0.802	−0.044
$Fe(H_2O)_6^{3+}$	2.0	−0.453	−0.340	−0.634	−0.204
$Fe(H_2O)_6^{2+}$	9.5	−0.660	−0.158	−0.789	−0.051
$Ga(H_2O)_6^{3+}$	2.6	−0.446	−0.343	−0.632	−0.199
$Hf(H_2O)_6^{4+}$	0.2	−0.268	−0.509	−0.516	−0.317
$In(H_2O)_6^{3+}$	4.0	−0.465	−0.329	−0.637	−0.196
$K(H_2O)_6^{+}$	14.5	−0.865	0.007	−0.928	0.069
$Li(H_2O)_6^{+}$	13.9	−0.853	0.003	−0.908	0.062
$Mg(H_2O)_6^{2+}$	11.4	−0.660	−0.162	−0.790	−0.054
$Mn(H_2O)_6^{2+}$	10.6	−0.664	−0.155	−0.790	−0.049
$Na(H_2O)_6^{+}$	14.7	−0.859	0.003	−0.932	0.074
$Ni(H_2O)_6^{2+}$	9.9	−0.655	−0.159	−0.783	−0.056
$Sc(H_2O)_6^{3+}$	4.3	−0.465	−0.330	−0.637	−0.202
$Sn(H_2O)_6^{4+}$	−0.6	−0.254	−0.517	−0.510	−0.316
$Sr(H_2O)_6^{2+}$	13.3	−0.689	−0.139	−0.796	−0.051
$Ti(H_2O)_6^{3+}$	2.2	−0.457	−0.341	−0.635	−0.205
$Ti(H_2O)_6^{4+}$	−4.0	−0.243	−0.535	−0.509	−0.327
$Tl(H_2O)_6^{+}$	13.2	−0.866	0.007	−0.888	0.028
$Tl(H_2O)_4^{3+}$	0.6	−0.411	−0.365	−0.641	−0.190
$Zn(H_2O)_6^{2+}$	9.5	−0.660	−0.158	−0.790	−0.051
$Zr(H_2O)_6^{4+}$	−0.3	−0.272	−0.508	−0.518	−0.317

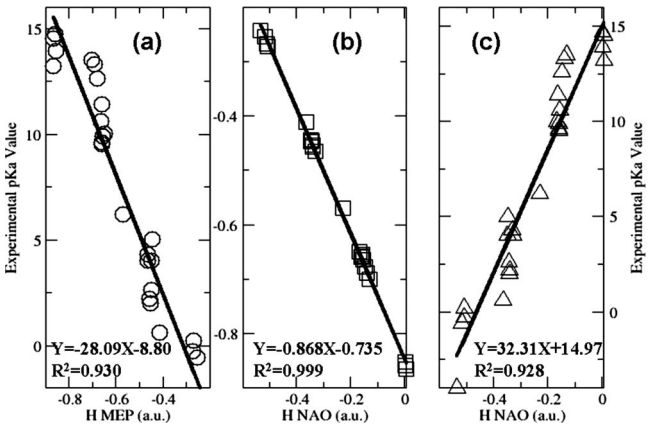


FIG. 4. Linear relationships between MEP at hydrogen nucleus (H MEP), natural valence atomic orbital energy of hydrogen (H NAO), and experimental pKa values for the 27 $M(H_2O)_n^{m+}$ complexes: (a) H MEP vs experimental pKa, (b) H NAO vs H MEP, and (c) H NAO vs experimental pKa data.

rating that molecular acidity is a local property limited to the neighborhood of the acidic atom and that the impact of the environment, through space or through bonds, is reflected through the stereoelectronic property change near the neighboring region. Figures 4(b) and 4(c) are the reminiscent of the similar relationship for the acidic atom in our earlier work,[14] revealing that MEP_H and NAO_H are equivalent as a molecular acidity descriptor. This result illustrates that the electrostatic potential on the nucleus of the dissociating hydrogen results predominantly from the contribution of the valence NAO of the same atom. For hydrogen atoms, this assertion is certainly valid because (i) they do not have any core electrons and (ii) they can be considered partially charged in the molecule and no electrons from neighboring atoms contribute to MEP_H.

V. CONCLUSIONS

Extending our earlier work on predicting molecular acidity with the nuclear electrostatic potential and valence NAO energies, in this work we establish the relationships on the basis of conceptual DFT. We show that as a first-order approximation, the pKa value of an acid should be proportional to the electrostatic potential on the nucleus of the dissociating proton. With 27 main groups and first to third row transition metal-water complexes as illustrative examples, the validity of this linear relationship has been confirmed. We also verified the equivalence between the nuclear electrostatic potential and valence NAO energy as the molecular acidity descriptor for both acidic atom, oxygen in this case, and the leaving proton. In addition, results from both our method and the conventional approach through the thermodynamic cycle have shown that the methods are similar in accuracy.

ACKNOWLEDGMENTS

Helpful discussions with Robert G. Parr of University of North Carolina and Andy Maynard are gratefully acknowledged. This work was supported in part by the National Institutes of Health (Grant No. HL-06350), the NSF (FRG Grant No. DMR-0804549), and the Intramural Research Program of NIH, NIEHS. We acknowledge the use of the computational resources provided by the Research Computing Center at University of North Carolina at Chapel Hill and the Biomedical Unit of the Pittsburgh Supercomputer Center.

[1] W. L. Jorgensen, J. M. Briggs, and J. Gao, J. Am. Chem. Soc. **109**, 6857 (1987).
[2] C. Lim, D. Bashford, and M. Karplus, J. Phys. Chem. **95**, 5610 (1991).
[3] M. J. Potter, M. K. Gilson, and J. A. McCammon, J. Am. Chem. Soc. **116**, 10298 (1994); J. E. Nielsen and J. A. McCammon, Protein Sci. **12**, 1894 (2003); C. A. Fitch and B. Garcia-Moreno, Biophys. J. **86**, 86A (2004).
[4] G. Alagona, C. Ghio, and P. A. Kollman, J. Am. Chem. Soc. **117**, 9855 (1995).
[5] P. Politzer, Theor. Chem. Acc. **111**, 395 (2004); P. Politzer, Y. G. Ma, A. F. Jalbout, and J. S. Murray, Mol. Phys. **103**, 15 (2005).
[6] G. da Silva, E. M. Kennedy, and B. Z. Dlugogorski, J. Phys. Chem. A **110**, 11371 (2006); K. Ohno and M. Sakurai, J. Comput. Chem. **27**, 906 (2006); C. M. MacDermaid and G. A. Kaminski, J. Phys. Chem. B **111**, 9036 (2007).
[7] A. Klamt, F. Eckert, M. Diedenhofen, and M. E. Beck, J. Phys. Chem. A **107**, 9380 (2003); M. Namazian and S. Halvani, J. Chem. Thermodyn. **38**, 1495 (2006).
[8] H. Li, A. D. Robertson, and J. H. Jensen, Proteins: Struct., Funct., Bioinf. **4**, 704 (2007); N. Goudarzi and M. Goudarzi, Mol. Phys. **107**, 1495 (2009).
[9] J. E. Nielsen, Methods Enzymol. **454**, 233 (2009); J. C. Dearden, M. T. D. Cronin, and D. C. Lappin, J. Pharm. Pharmacol. **59**, A7 (2007); A. P. Harding, D. C. Wedge, and P. L. A. Popelier, J. Chem. Inf. Model. **49**, 1914 (2009).
[10] F. De Proft, S. Amira, K. Choho, and P. Geerlings, J. Phys. Chem. B **98**, 5227 (1995).
[11] H. J. S. Machado and A. Hinchliffe, J. Mol. Struct.: THEOCHEM **339**, 255 (1995).
[12] R. C. Deka, R. K. Roy, and K. Hirao, Chem. Phys. Lett. **389**, 186 (2004).
[13] K. Gupta, D. R. Roy, V. Subramanian, and P. K. Chattaraj, J. Mol. Struct.: THEOCHEM **812**, 13 (2007).
[14] S. B. Liu and L. G. Pedersen, J. Phys. Chem. A **113**, 3648 (2009).
[15] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
[16] P. Geerlings, F. De Proft, and W. Langenaeker, Chem. Rev. (Washington, D.C.) **103**, 1793 (2003); P. Geerlings and F. De Proft, Phys. Chem. Chem. Phys. **10**, 3028 (2008).
[17] P. K. Chattaraj, U. Sarkar, and D. R. Roy, Chem. Rev. (Washington, D.C.) **106**, 2065 (2006).
[18] S. B. Liu, Acta Phys.-Chim. Sin. **25**, 590 (2009).
[19] J. R. Pliego, Chem. Phys. Lett. **367**, 145 (2003).
[20] S. B. Liu, T. Li, and P. W. Ayers, J. Chem. Phys. **131**, 114106 (2009).
[21] P. W. Ayers, S. B. Liu, and T. Li, Chem. Phys. Lett. **480**, 348 (2009).
[22] G. Wulfsberg, *Principles of Descriptive Inorganic Chemistry* (University Science Books, Sausalito, CA, 1991), p. 25.
[23] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[24] A. D. Becke, J. Chem. Phys. **104**, 1040 (1996).
[25] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).
[26] P. J. Hay and W. R. Wadt, J. Chem. Phys. **82**, 270 (1985).
[27] S. B. Liu and Y. X. Yu, J. Mol. Struct.: THEOCHEM **235**, 115 (1991).
[28] S. B. Liu and W. Langenaeker, Theor. Chem. Acc. **110**, 338 (2003).
[29] A. G. Zhong and S. B. Liu, J. Theor. Comput. Chem. **4**, 833 (2005).
[30] C. Rong, S. Lian, D. Yin, B. Shen, A. Zhong, L. Bartolotti, and S. B. Liu, J. Chem. Phys. **125**, 174102 (2006).
[31] J. K. Badenhoop and F. Weinhold, J. Chem. Phys. **107**, 5406 (1997).
[32] V. Barone and M. Cossi, J. Phys. Chem. A **102**, 1995 (1998).
[33] M. Cossi, N. Rega, G. Scalmani, and V. Barone, J. Comput. Chem. **24**, 669 (2003).
[34] M. Liptak and G. C. Shields, Int. J. Quantum Chem. **85**, 727 (2001); J. Am. Chem. Soc. **123**, 7314 (2001).
[35] J. Y. Lee, B. J. Byun, and Y. K. Kang, J. Phys. Chem. B **112**, 11189 (2008).
[36] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford CT, 2004.
[37] R. G. Parr and R. G. Pearson, J. Am. Chem. Soc. **105**, 7512 (1983).
[38] R. G. Pearson, J. Chem. Educ. **64**, 561 (1987).

[39] R. G. Pearson, *Chemical Hardness* (Wiley-VCH, Weinheim, 1997).
[40] P. W. Ayers, Faraday Discuss. **135**, 161 (2007).
[41] R. G. Parr, L. von Szentpaly, and S. B. Liu, J. Am. Chem. Soc. **121**, 1922 (1999).
[42] S. B. Liu, in *Chemical Reactivity Theory: A Density Functional View*, edited by P. K. Chattaraj (Taylor & Francis, Boca Raton, 2009), p. 179.
[43] A. Cedillo, R. Contreras, M. Galvan, A. Aizman, J. Andres, and V. S. Safont, J. Phys. Chem. A **111**, 2442 (2007).
[44] J. J. Klicic, R. A. Friesner, S.-Y. Liu, and W. C. Guida, J. Phys. Chem. A **106**, 1327 (2002).